

WHAT IS CLAIMED IS:

1. A method for managing HPC node failure comprising:

5 determining that one of a plurality of HPC nodes has failed, each HPC node comprising an integrated fabric; and

removing the failed node from a virtual list of HPC nodes, the virtual list comprising one logical entry for each of the plurality of HPC nodes.

10

2. The method of Claim 1, further comprising:

determining that at least a portion of an HPC job was being executed on the failed node; and terminating the HPC job.

15

3. The method of Claim 2, further comprising:

determining that the HPC job was associated with a subset of the plurality of HPC nodes; and deallocating the subset of HPC nodes.

20

4. The method of Claim 3, each entry of the virtual list comprising a node status and the method further comprising changing the status of each of the subset of HPC nodes to "available".

25

5. The method of Claim 3, further comprising:
determining dimensions of the terminated job based
on one or more job parameters and an associated policy;
dynamically allocating a second subset of the
5 plurality of HPC nodes based on the determined
dimensions; and
executing the terminated job on the allocated second
subset.

10 6. The method of Claim 5, the second subset
comprising a substantially similar set of nodes to the
first subset.

7. The method of Claim 5, wherein dynamically
15 allocating the second subset comprises:
determining an optimum subset of nodes from a
topology of unallocated HPC nodes; and
allocating the optimum subset.

20 8. The method of Claim 1, further comprising:
locating a replacement HPC node for the failed HPC
node; and
updating the logical entry of the failed HPC node
with information on the replacement HPC node.

25 9. The method of Claim 1, wherein determining one
of the plurality of HPC nodes has failed comprises
determining that a repeating communication has not been
received from the failed node.

30

10. The method of Claim 1, wherein determining one of the plurality of HPC nodes has failed is accomplished through polling.

5 11. Software for managing HPC node failure operable to:

determine that one of a plurality of HPC nodes has failed, each node comprising an integrated fabric; and

10 remove the failed node from a virtual list of HPC nodes, the virtual list comprising one logical entry for each of the plurality of HPC nodes.

12. The software of Claim 11, further operable to:
determine that at least a portion of an HPC job was
15 being executed on the failed node; and
terminate the HPC job.

13. The software of Claim 12, further operable to:
determine that the HPC job was associated with a
20 subset of the plurality of HPC nodes; and
deallocate the subset of HPC nodes.

14. The software of Claim 13, each entry of the virtual list comprising a node status and the software
25 further operable to change the status of each of the subset of HPC nodes to "available".

15. The software of Claim 13, further operable to:
determine dimensions of the terminated job based on
one or more job parameters and an associated policy;

5 dynamically allocate a second subset of the
plurality of HPC nodes based on the determined
dimensions; and

execute the terminated job on the allocated second
subset.

10 16. The software of Claim 15, the second subset
comprising a substantially similar set of nodes to the
first subset.

17. The software of Claim 15, wherein the software
15 operable to dynamically allocate the second subset
comprises software operable to:

determine an optimum subset of nodes from a topology
of unallocated HPC nodes; and

allocate the optimum subset.

20

18. The software of Claim 11, further operable to:

locate a replacement HPC node for the failed HPC
node; and

25 update the logical entry of the failed HPC node with
information on the replacement HPC node.

19. The software of Claim 11, wherein the software
operable to determine one of the plurality of HPC nodes
has failed comprises software operable to determine that
30 a repeating communication has not been received from the
failed node.

20. The software of Claim 11, wherein the software operable to determine one of the plurality of HPC nodes has failed is accomplished through polling.

5 21. A system for managing HPC node failure comprising:

 a plurality of HPC nodes, each node including an integrated fabric; and

 a management node operable to:

10 determine that one of the plurality of HPC nodes has failed, each node comprising an integrated fabric; and

 remove the failed node from a virtual list of HPC nodes, the virtual list comprising one logical entry
15 for each of the plurality of HPC nodes.

22. The system of Claim 21, the management node further operable to:

 determine that at least a portion of an HPC job was
20 being executed on the failed node; and
 terminate the HPC job.

23. The system of Claim 22, the management node further operable to:

25 determine that the HPC job was associated with a subset of the plurality of HPC nodes; and
 deallocate the subset of HPC nodes.

24. The system of Claim 23, each entry of the virtual list comprising a node status and the management node further operable to change the status of each of the subset of HPC nodes to "available".

5

25. The system of Claim 23, the management node further operable to:

determine dimensions of the terminated job based on one or more job parameters and an associated policy;

10 dynamically allocate a second subset of the plurality of HPC nodes based on the determined dimensions; and

execute the terminated job on the allocated second subset.

15

26. The system of Claim 25, the second subset comprising a substantially similar set of nodes to the first subset.

20 27. The system of Claim 25, wherein the management node operable to dynamically allocate the second subset comprises the management node operable to:

determine an optimum subset of nodes from a topology of unallocated HPC nodes; and

25 allocate the optimum subset.

28. The system of Claim 21, the management node further operable to:

locate a replacement HPC node for the failed HPC node; and

5 update the logical entry of the failed HPC node with information on the replacement HPC node.

29. The system of Claim 21, wherein the management node operable to determine one of the plurality of HPC
10 nodes has failed comprises the management node operable to determine that a repeating communication has not been received from the failed node.

30. The system of Claim 21, wherein the management
15 node operable to determine one of the plurality of HPC nodes has failed is accomplished through polling.